**International Conference on Applied Energy 2019**
**Aug 12-15, 2019, Västerås, Sweden**
**Paper ID: 200**

# REINFORCEMENT LEARNING–BASED ENERGY MANAGEMENT STRATEGY FOR A SERIES-PARALLEL HYBRID BUS

Han Xuefeng[1,2], He Hongwen[1,*], Wu Jingda[1], Peng Jiankun[1] , Chen Rui[3,2]
1 School of Mechanical Engineering, Beijing Institute of Technology
2 China North Vehicle Research Institute
3 State Key Laboratory of Automotive Safety and Energy, Tsinghua University

**ABSTRACT**

An energy management strategy based on double deep Q-learning algorithm is proposed for a Series-Parallel Hybrid Bus. The models of powertrain configuration and its main components are first established. Subsequently, a rule-based energy management strategy will be proposed. The China typical urban driving cycle (CTUDC) is used to evaluate the fuel economy performance of the two strategies studied in this paper. The simulation result indicates that the energy management strategy based on reinforcement learning decreased the fuel consumption by 7.3% per 100km compared to rule-based strategy.

**Keywords:** Series-Parallel Hybrid Bus; energy management; rule-based; double deep Q-learning

## 1. INTRODUCTION

Air quality has become a serious concern in cities and urban areas in recent years. Environmental concerns and the security of the energy supply have pushed Chinese government to explore environmentally friendly, efficient and sustainable transportation solution[1]. In response to these crises, Hybrid electric vehicles (HEV) / Plug-in hybrid electric vehicles (PHEV) have been proved an effective solution for the energy and environmental problems.

The widely accepted HEV configurations can be classified into three categories: series, parallel and power-split [2]. The series-parallel configuration HEV combines the advantages of both series and parallel configurations and has been adopted by an increasing number of bus OEMs (such as Yutong and Higer) in China[1]. As a result of multiple power sources, HEVs have more degrees of freedom to supply the power demand, compared with the conventional vehicles[3]. Therefore, the energy management strategy targeting to maximize the overall powertrain efficiency and minimize fuel consumption will becoming the research focus of the Serial–Parallel Hybrid Electric Bus.

Energy management strategies for HEV can be generally classified into two typical types: rule-based and optimization-based. Rule-based energy management strategies have been widely used due to their reliability, but they are all sub-optimal and highly dependent on engineering experience. Charge depleting–charge sustaining (CD–CS) mode is one of the typical instances[3]. The optimization-based strategies adjust the control variables by minimizing the predefined cost function under feasible constraints. Typical energy management strategies based on optimization include Equivalent Consumption Minimization Strategy (ECMS), model predictive control (MPC), dynamic programming algorithm (DP) [4,5], etc. In general, the strategy based on optimization requires a large amount of computation and is difficult to guarantee real-time performance.

This paper provides a powertrain configuration of a Series-Parallel Hybrid Electric Bus, and a rule-based energy management strategy for the bus will be proposed. Then, a new strategy based on reinforcement learning will be used to improve the performance.

The reminder of this paper is organize as follows: In section 2, the bus configuration model is introduced. The rule-based strategy is developed in section 3 and an improved energy management strategy based on reinforcement learning is presented in section 4. The simulation results are shown in section 5. Finally, comments and conclusions are discussed.

## 2. SERIAL-PARALLEL HYBRID BUS MODEL

### 2.1 System configuration

The powertrain configuration of the Series-Parallel Hybrid Electric Bus studied in this paper is shown in figure 1, which consists of a Compressed Natural Gas (CNG) engine, an integrated starting generator (ISG), a traction motor, a battery pack, and an electro-controlled clutch which be used to implement the switch between serial and parallel modes.

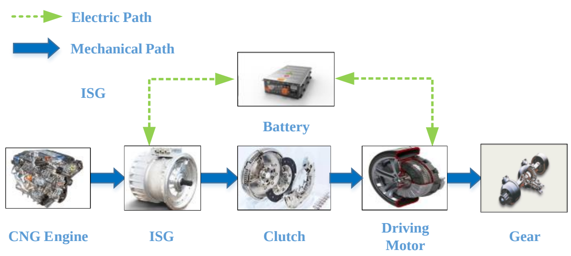


Fig. 1 Architecture of the Series-Parallel Hybrid Electric Bus

### 2.2 Power Demand Model

When the driving cycle is known a priori, the power demand $P_{req}$ can be defined as:

$$\begin{cases} P_{req} = (F_r + F_i + F_a)\bar{v} \\ F_r = mgf_r \\ F_i = ma \\ F_a = (C_d A/21.15)\bar{v}^2 \end{cases} \quad (1)$$

Where $F_r$ is the rolling resistance, $m$ is the total mass of the vehicle, $g$ is the gravity coefficient, and $f_r$ is the coefficient of rolling resistance. $F_i$ is the inertial force, $a$ is the acceleration. $F_a$ is the aerodynamic drag, $C_d$ is the aerodynamic coefficient, and $A$ is the fronted area. $\bar{v}$ is the vehicle speed

### 2.3 Powertrain model

The experimental approach is adopted to model the engine, ignoring the dynamic characteristics of the engine, the quasi-static model of the engine is established. Therefore, engine fuel consumption is only related to two parameters: current engine speed and actual engine torque. The engine fuel consumption map is expressed as a non-linear 3-D MAP in fig.2.

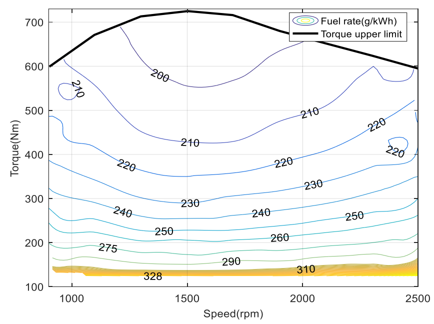


Fig. 2 Engine map

The powertrain configuration in this research choose a 510V Li-ion battery as the electricity sources. The state of charge (SOC) in the battery is chosen as the state variable. A simple and effective internal resistance battery model is adopted to describe the basic dynamics as follows:

$$\frac{dSOC}{dt} = -\frac{V_{oc} - \sqrt{V_{oc} - 4R_{\text{int}} \cdot P_{bat}}}{2R_{\text{int}} Q_{bat}} \quad (3)$$

Which $V_{oc}$ denotes the open-circuit voltage, $R_{\text{int}}$ denotes the internal resistance, and $P_{bat}$ denotes the battery load power. All of them are a function of SOC.

## 3. RULE-BASED ENERGY MANAGEMENT APPROACH

Rule-based energy management strategy is widely used due to its reliability. Furthermore, the rule-based energy management is simple to implement in real-time control system. In this paper, the Series-Parallel Hybrid Bus work mode division depends on battery SOC, power requirements and vehicle Velocity. The work condition rules are simplified as figure 3a and figure 3b.

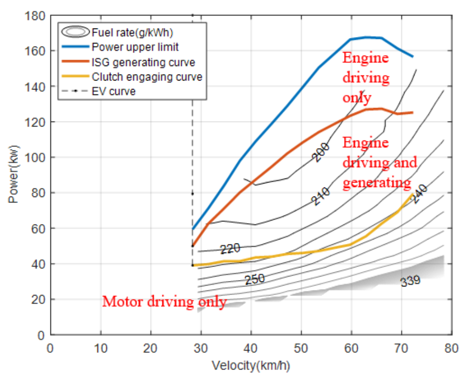


Fig. 3a Working mode with SOC below 65%

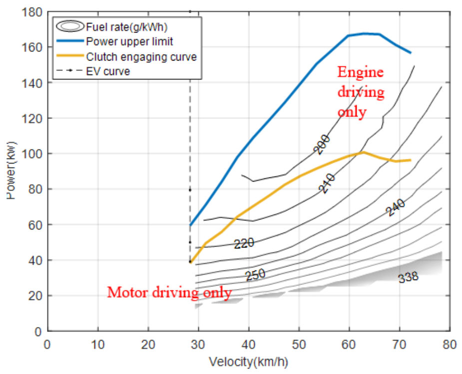


Fig. 3b Working mode with SOC above 65%

When SOC < 65%, The ISG and engine is shutdown when vehicle power demand is under the clutch engaging curve, so the efficiency loss caused by the growth of energy circulation can be avoided. In this Series mode, the clutch is disengaged and Bus is driving by motor only. Further, when the vehicle power demand is above the clutch engaging curve and Bus drives at high velocity, the clutch will be engaged, and the Bus will be working in the Parallel mode. Considering the efficiency of the engine, some of the engine power will be used for driving, and the rest will be used for ISG motor generation. When the demand power of the vehicle is larger than the ISG power generation curve, the engine will work in the high efficient zone, and Bus is driving by engine only. In addition to all above, when the vehicle speed is less than 27km/h, the Bus is in EV mode with only the motor driving.

When SOC > 65%, and the driving demand is satisfied by the traction motor, the energy will be provided by only the power battery pack. In this mode, the regenerative braking is not permitted to avoid the possible over charging due to the higher SOC and the lower charge efficiency. Furthermore, when the power demand of the vehicle is greater than the power of the traction motor in the high-efficiency zone, the clutch will be engaged and the Bus will enter parallel mode. Bus will be driven by engine only.

It is worth mentioning that the power system configuration involved in this paper has a working mode that the bus is driven by a combination of engine, ISG and drive motor. This operation mode usually occurs when the Bus is running at high speed for rapid acceleration. This kind of operation condition is rarely seen for city buses, so it will be not discussed in this paper.

## 4. ENERGY MANAGEMENT APPROACH BASED ON DOUBLE DEEP Q LEARNING

For the energy management problem of hybrid electric vehicles, the demand power of vehicle driving cycle changes randomly, which can be considered as Markov decision processes(MDP). According to literature [4], the driving schedule can be considered a finite MDP. In this article, the driving schedule will be redefined as:

$$S_t = \{SOC(t), P_{\text{req}}(t), \bar{v}(t), a(t)\} \tag{4}$$

$$A_t = \{s_{clutch}(t), n_{isg}(t), T_e(t), T_m(t) \\ \left| s_{clutch} \in [0,1], n_g \in [0,2500], T_e \in [0,710], T_m[0,2100]\right\} \tag{5}$$

$$R_t = f(fuel(t), SOC(t)) \tag{6}$$

$S_t$ represents the set of state variables for each step. $A_t$ is the set of actions for each step. $s_{clutch}$ is the state of clutch, furthermore, $s_{clutch}=0$ represents the clutch is engaged, while $s_{clutch}=1$ means that the clutch is disengaged, $n_{isg}$ is the speed of ISG, $T_e$ is the torque of CNG engine, $T_m$ is the torque of driving motor. $R_t$ represents the reward for each actions taking, which is a function of SOC and CNG engine fuel consumption $fuel(t)$.

The reinforcement learning algorithm can effectively improve the control performance for Markov decision problems. The basic idea of reinforcement learning is creating a learning agent which can be able to sense the state of environment and to take actions that affect the state. The cumulative reward signal of learning agent can be maximized through the trial-and-error based on the analysis of actions that the agent takes [5].

Double deep Q-learning (DDQL) is the modification algorithm of a classical reinforcement learning algorithm which is called deep Q-learning(DQL). Standard DQL derives the optimal control strategy $\pi$ by maximizing the mathematical expectation of the total reward $\sum R_t$ .Consequently, the optimal value function $Q^*$ which guides the decision process of policy can be defined as distribution over the given current state $S_t$ and control action $A_t$ :

$$Y_t = R_{t+1} + \lambda \max_A Q(S_{t+1}, A) \quad (8)$$

Therefore, the iterative update process can be expressed as:

$$\begin{aligned} &Q(S_t, A_t; \theta_t) \leftarrow Q(S_t, A_t; \theta_t) + \\ &\alpha(Y_t - Q(S_t, A_t; \theta_t))\nabla_{\theta_t}(Q(S_t, A_t; \theta_t)) \end{aligned} \quad (9)$$

where α ∈ [0,1] means the decaying factor influencing the learning rate.

In the DDQL algorithm, two value functions $Q(S,A;\theta)$ and $Q'(S,A;\theta')$ with two different sets of weights $\theta$ and $\theta'$ are used to decouple the action selection from the reward evaluation in the formula (8) and (9), which can avoid the overoptimistic value estimates in the standard DQL algorithm resulting by the selections of overestimated values[6]. Thus, target values and the iteration process of DDQL can be redefined as:

$$Y_t = R_{t+1} + \lambda Q^{'}(S_{t+1}, \arg\max_A Q(S_{t+1}, A; \theta_t); \theta_t^{'}) \quad (10)$$

$$\begin{aligned} &Q(S_t, A_t; \theta_t) \leftarrow Q(S_t, A_t; \theta_t) + \\ &\alpha(Y_t - Q(S_t, A_t; \theta_t))\nabla_{\theta_t}(Q(S_t, A_t; \theta_t)) \end{aligned} \quad (11)$$

## 5. RESULTS AND DISCUSSION

In this section, the results obtained from the aforementioned rule-based strategy and DDQL are compared and discussed. The China typical urban driving cycle (CTUDC) is used for the simulation to evaluate the fuel economy. Simulation results in figure 7 show that both rule-based and ddql-based energy management strategies can maintain battery power within the set range. In order to compare the fuel consumption optimizations of the two algorithms, this paper will adopt the conversion method of battery power to fuel consumption proposed in literature [7].

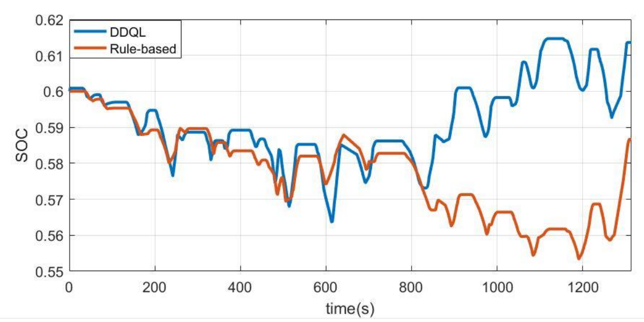


Fig. 4 SOC trajectories in Rule-based and DDQL

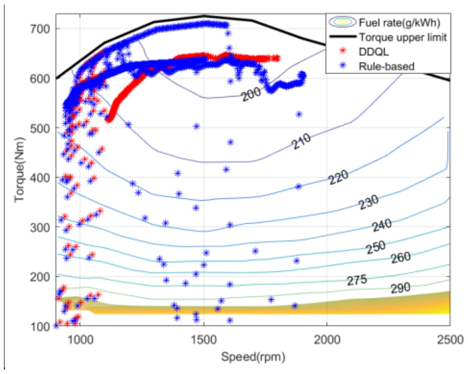


Fig. 5 Engine points in Rule-based and DDQL algorithms

The engine work point distributions of rule-based and DDQL strategy are shown in figure 5. Although the rule-based strategy can effectively control more engine operating points in the efficient area, the ddql-based strategy can further optimize the distribution of operating points to achieve better fuel economy.

Finally, the accurate fuel consumption statistics compensated by SOC-correction method are shown in table 2. The results indicates that, for the Series-Parallel Hybrid Electric Bus studied in this paper, the economic performance is much better than the conventional bus with same class. Furthermore, energy management strategy based on DDQL can achieve better fuel consumption. The fuel consumption result is 7.3% lower than the rule-based algorithm.

TABLE. 2 COMPARISON OF FUEL CONSUMPTION

| Algorithm | Fuel consumption (L/100km) | Fuel economy(%) |
|---|---|---|
| Rule-based | 20.99 | 61.6 |
| DDQL | 18.50 | 54.3 |
| **Conventional**[8] | **34.10** | **100** |

## 6. CONCLUSION

In this paper, a Series-Parallel Hybrid Electric Bus is introduced and its powertrain configuration model is proposed. Subsequently, rule-based and DDQL-based energy management algorithms will be proposed respectively. The performance of the proposed strategies are evaluated by using CTUDC. From simulation results indicate that DDQL-based Energy management strategy achieves significant fuel economy improvement compared with rule-based strategy, the fuel consumption decreased from 20.99L/100km to 18.5L/100km.

**ACKNOWLEDGEMENT**

This project is supported by National Key R&D Program of China (Grand No.2018YFB0105900) in part, and the National Natural Science Foundation of China (Grand No.51675042&51705020） in Part.

**REFERENCE**

[1] Yuanchun Cai. Minggao Ouyang. Fuyuan Yang. Energy Management and Design Optimization for A Series-Parallel PHEV City Bus. International Journal of Automotive Technology, Vol. 18, No. 3, pp.473−487. 2017.

[2] M. Sabri, M.F.，Danapalasingam, K.A.，Rahmat, M.F. A review on hybrid electric vehicles architecture and energy management strategies [J]. Renewable & Sustainable Energy Reviews，2016，53：1433-1442.

[3]J. Gonder. T. Markel, Energy management strategies for plug-in hybrid electric vehicles, Soc. Automotive Eng., Int., Warrendale, PA, USA, SAE Tech, 2017.

[4] Julien Lescot. Antonio Sciarretta. On the integration of optimal energy management and thermal management of hybrid electric vehicles. 2010 IEEE Vehicle Power and Propulsion Conference. IEEE, 2010.

[4] C.M, Martinez. X.S, Hu. D.P, Cao. E. Velenis. B, Gao. M. Wellers. Energy Management in Plug-in Hybrid Electric Vehicles: Recent Progress and a Connected Vehicles Perspective. IEEE Transactions on Vehicular Technology, Vol. 66, NO.6, June 2017.

[5] Mojtaba Shams-Zahraei, Abbas Z. Kouzani, Steffen Kutter, Bernard Bäker. Integrated thermal and energy management of plug-in hybrid electric vehicles. Journal of power sources 216 (2012): 237-248.

[6] Brazdil, T. Chatterjee, K. Chmelik, M. Forejt, V. Kretinsky, J. Kwiatkowsha, M. Parker, D. Ujma, M. Verification of markov decision processes using learning algorithms. Logic Comput. Sci. 2014, 2, 4–18.

[7] Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An Introduction, 2th, MIT Press., 2017.

[8] Hado van Hasselt, Arthur Guez, David Silver, Deep Reinforcement Learning with Double Q-learning, Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 2016.

[9] Zou, Z.Y. Xu, J. Mi, C. Cao, B.G. Evaluation of model based state of charge estimation methods for lithium-Ion batteries. Energies 2014, 7, 5065–5082.

[10] Fuel consumption limits for heavy-duty commercial vehicles. GB 30510-2018 .ICS 43.020. T40.